INFORMATION AND CONTROL 28, 91–112 (1975)

Developmental Systems with Interaction and Fragmentation

K. RUOHONEN

Department of Mathematics, University of Turku, 20500 Turku 50, Finland

Communicated by Arto Salomaa

A new class of developmental systems with interaction is introduced. These systems are capable of splitting words into two or more separate parts. This property can be used to model reproduction processes and discarding parts of an organism. It is also of interest from the formal language theory point of view. This paper discusses the basic properties of the families of languages obtained, as well as their position with respect to other families of developmental languages and the families of the Chomsky hierarchy.

1. INTRODUCTION

In 1968 A. Lindenmayer introduced mathematical models (now customarily called L systems) for the development of hypothetical one-dimensional organisms. They were originally linear arrays of automata. With suitable encodings these models have been successfully used to simulate the development of certain living organisms or parts of them. The difference between the von Neumann type tessellation automata and L systems is that the latter allow cell division and death at any place in the organism, not only at the ends. Later the formalism of L systems was changed to that of a generative grammar with parallel rewriting. This gave rise to various families of languages (developmental languages) the properties of which differ considerably from those of the families studied in classical formal language theory.

We investigate the effect of adding the following novel property (introduced in Rozenberg, Ruohonen and Salomaa, 1974, for L systems without interaction) to L systems. In addition to ordinary rewriting a symbol may have the ability to induce a cut in the scanned string. We call these L *systems with fragmentation* (referred to as JL systems). Thus (excluding the effects of possible nondeterminacy) by a JL system a string can directly generate a finite set of strings (a clone, in biological terms) instead of just one string.

This property can be used to simulate reproduction processes, such as the production of gametes or spores and vegetative reproduction through fragmentation (the propagation of strawberry by runners, to pick a familiar example). It also provides us with a new formalism of cell death. (See Rozenberg, Ruohonen and Salomaa (1974) for a more detailed discussion about the biological significance of JL systems.)

We give briefly the contents of each section.

In Section 2 we give the definitions of JL systems and JL languages.

In Section 3 we discuss some of the properties of JL languages. Several normal forms are given for JL systems and the hierarchy of JL languages, with respect to the amount of context available, is established. Also an alternative way of defining JL languages (using "ordinary" L systems and a special operator) is given.

In Section 4 we introduce a special form of a JL system, a JL system with *erasing fragmentation* (or a jL system), in which the fragmentation process always involves cell death. The properties of jL languages are discussed and the generating capacity of jL systems is compared with that of JL systems and IL systems.

In Section 5 it is shown that all the families of JL languages and jL languages investigated are anti-AFL's.

In Section 6 the generating capacity of JL systems and jL systems is compared with that of various developmental systems and the grammars of the Chomsky hierarchy.

2. Preliminaries and Definitions

We expect the reader to be acquainted with the basics and standard notions of formal language theory (see, e.g., Salomaa, 1973) and developmental systems theory (see, e.g., Herman and Rozenberg, 1974). The following notations need to be specified: by an alphabet we mean a nonempty finite set; the empty word is denoted by λ; (strict) inclusion is denoted by $\subset$ ($\subsetneq$); the length of a word P is denoted by $\lg(P)$; throughout the paper, k and l are assumed to be arbitrary (but usually fixed within a certain context) nonnegative integers, unless other restrictions are stated.

Let A be an alphabet and q a symbol (possibly not in A). A word $P \in (A \setminus \{q\})^*$ is a *q-guarded subword* of a word $Q \in A^*$ iff qPq is a subword of qQq. We define the operator J_q by the equation

$$J_q(L) = \{P \mid P \text{ is a } q\text{-guarded subword of some word of } L\} \quad (\text{for } L \subset A^*).$$

We note in passing the following (later not explicitly used) properties (i)–(iv) of the operator J_q.

(i) $J_q(L) = L$ for any language L over an alphabet not containing q.

(ii) For any symbols q and q', the operators J_q and $J_{q'}$ commute.

(iii) $J_q(L) \neq \Phi$ for any language $L \neq \Phi$.

(iv) $J_q(L) \subset (A \setminus \{q\})^*$ for any language L over the alphabet A. Thus J_q is idempotent.

A Lindenmayer system with $\langle k, l\rangle$ interaction and fragmentation (*a* $J\langle k, l\rangle L$ system in short) is an ordered quadruple (A, δ, g, ω) where A is an alphabet, g is a symbol not in A, ω is a word over $A_g = A \cup \{g\}$ (the *axiom*) and δ (the *production relation*, the elements of which are called *productions*) is defined as follows: Let $A_G = B_G \times A \times C_G$ where

$$B_G = \bigcup_{n=0}^{k} \{g^n\} A^{k-n}, \qquad C_G = \bigcup_{n=0}^{l} A^{l-n}\{g^n\}.$$

Then δ is a finite subset of $A_G \times A_g{}^*$ such that, for each $X \in A_G$, $(X, P) \in \delta$ for some $P \in A_g{}^*$.

Instead of $((P, a, Q), R) \in \delta$ we often write $(P, a, Q) \rightarrow_\delta R$ (also this expression is called a production). A production $X \rightarrow_\delta P$ is called *deterministic* iff $\delta \cap (\{X\} \times A_g{}^*) = \{(X, P)\}$, and *erasing* iff $P = \lambda$. A $J\langle k, l\rangle L$ system is called *deterministic* (or a $DJ\langle k, l\rangle L$ system) iff all its productions are deterministic (i.e., its production relation is a function), and *propagating* (or a $PJ\langle k, l\rangle L$ system) iff none of its productions is erasing.

Let $G = (A, \delta, g, \omega)$ be a $J\langle k, l\rangle L$ system. We define the relation $\Rightarrow_G$ on the set $A^* \times A^*$ as follows. First $(\lambda, \lambda) \in \Rightarrow_G$. Let $P = a_1 \cdots a_m$ where $m \geqslant 1$ and $a_1, \ldots, a_m \in A$. Furthermore let L_i and R_i be such words of B_G and C_G respectively that $L_i a_i R_i$ is a subword of $g^k P g^l$ (for $i = 1, \ldots, m$). Then, for each word $Q \in A^*$, $(P, Q) \in \Rightarrow_G$ iff there exists a word $R = S_1 \cdots S_m \in A_g{}^*$ such that

$$(L_i, a_i, R_i) \underset{\delta}{\rightarrow} S_i \qquad (\text{for } i = 1, \ldots, m)$$

and Q is a g-guarded subword of R. The reflexive-transitive closure of $\Rightarrow_G$ is denoted by $\overset{*}{\Rightarrow}_G$. For $(P, Q) \in \Rightarrow_G$ we write also $P \Rightarrow_G Q$ (a *direct derivation*) and say that P *G-generates Q directly*. Correspondingly, for $(P, Q) \in \overset{*}{\Rightarrow}_G$ we write also $P \overset{*}{\Rightarrow}_G Q$ (a *derivation*) and say that P *G-generates Q*. The *language generated by G* is defined by

$$L(G) = \{Q \mid P \underset{G}{\overset{*}{\Rightarrow}} Q \text{ for some } P \in J_g(\{\omega\})\}.$$

Note that if, for a $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$, $\delta \subset A_G \times A^*$, then G is an $F\langle k, l\rangle L$ system. If in addition $\omega \in A^*$, then G is a $\langle k, l\rangle L$ system.

Following the customary notation, a $J\langle 0, 0\rangle L$ system is also called a $J0L$ system.

Finally we define the families

$$J\langle k, l\rangle \mathscr{L} = \{L \mid L = L(G) \text{ for some } J\langle k, l\rangle L \text{ system } G\},$$

$J0\mathscr{L} = J\langle 0, 0\rangle \mathscr{L}$ and

$$JI\mathscr{L} = \bigcup_{k=0}^{\infty} \bigcup_{l=0}^{\infty} J\langle k, l\rangle \mathscr{L},$$

and call the elements of them $J\langle k, l\rangle L$ languages, $J0L$ languages and JIL languages respectively. The family of $J0L$ languages thus defined coincides with the family of $J0L$ languages defined in Rozenberg, Ruohonen and Salomaa (1974).

3. Properties of $J\langle k, l\rangle L$ Languages

The following two theorems are straightforward consequences of definitions of the previous section.

Theorem 3.1. *For any nonnegative integers k, l, k' and l' such that $k \leqslant k'$ and $l \leqslant l'$, $J\langle k, l\rangle \mathscr{L} \subset J\langle k', l'\rangle \mathscr{L}$.* ∎

Theorem 3.2. *Each $J\langle k, l\rangle L$ language can be generated by a $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ such that $\delta \subset A_G \times (A^*\{g\}A^* \cup A^*)$.*

Proof. Let $L = L(G)$ for some $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$. We transform G to another $J\langle k, l\rangle L$ system $G' = (A, \delta', g, \omega')$ as follows. First $\delta' \cap (A_{G'} \times (A^*\{g\}A^* \cup A^*)) = \delta \cap (A_G \times (A^*\{g\}A^* \cup A^*)$. For each $P = P_1 g P_2 \cdots g P_m$ such that $m \geqslant 3, P_1, \ldots, P_m \in A^*$ and $\delta \cap (A_G \times \{P\}) \neq \Phi$ (if any, otherwise $G' = G$) we catenate the axiom from the right by $Q_2 \cdots Q_{m-1}$ where

$$Q_i = \begin{cases} gP_i, & \text{if } P_i \in L \\ \lambda, & \text{if } P_i \notin L \end{cases} \qquad (\text{for } i = 2, \ldots, m-1).$$

At the same time we change the corresponding productions $(R, a, Q) \to_\delta P$ to $(R, a, Q) \to_{\delta'} P_1 g P_m$. Obviously $L(G') = L(G)$ and

$$\delta' \subset A_{G'} \times (A^*\{g\}A^* \cup A^*). \quad ∎$$

The process described in the previous proof is not effective in general. This is seen by the following lemma (the unbracketed claim), the essence of which is the well-known fact that the membership problem for Lindenmayer systems with interaction is not algorithmically solvable.

LEMMA 3.1. *There exist* $J\langle 1, 0\rangle L$ *systems* $G = (A, \delta, g, \omega)$ (*resp.* $G' = (A, \delta', g, \omega')$) *and distinct symbols* $a, b, c \in A$ *such that*

(i) *there is no algorithm for deciding whether or not* $A^*\{ba\}A^* \cap L(G) = \Phi$ (*resp.* $\{a\}A^* \cap L(G') = \Phi$);

(ii) $(b, a, \lambda) \rightarrow_\delta gcg$ (*resp.* $(g, a, \lambda) \rightarrow_{\delta'} cga$);

(iii) $\delta \setminus \{((b, a, \lambda), gcg)\} \subset A_G \times (A \setminus \{c\})^*$

$$(\textit{resp. } \delta' \setminus \{((g, a, \lambda), cga)\} \subset A_{G'} \times (A \setminus \{c\})^*).$$

There also exist $J\langle 0, 1\rangle L$ *systems with analogous properties.*

Proof. We prove only the unbracketed claim, the proof of the bracketed one being analogous. Let $(B, n, (\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n))$ be an instance of the Post correspondence problem. Our $J\langle 1, 0\rangle L$ system is now $G = (A, \delta, g, \omega)$ where A is the union of the disjoint alphabets B and

$$I = \{(i) \mid i = 1, \ldots, n\}, \qquad C_1 = \{x' \mid x \in B \cup I\},$$
$$C_2 = \{x'' \mid x \in B \cup I \cup \{b\}\}, \qquad C_3 = \{a, b, c, d, e, f, h, \omega\},$$

and δ is defined by the following schemes I–XIII.

I	$(g, \omega, \lambda) \underset{\delta}{\rightarrow} b(i)\, d\alpha_i f(i)\, e\beta_i ab$	(for $i = 1, \ldots, n$)
II	$(x, b, \lambda) \underset{\delta}{\rightarrow} b^2$	(for $x \in A_g$)
III	$\begin{cases}(x, d, \lambda) \underset{\delta}{\rightarrow} (i)\, d\alpha_i \\ (x, e, \lambda) \underset{\delta}{\rightarrow} (i)\, e\beta_i\end{cases}$	(for $x \in I$ and $i = 1, \ldots, n$)
IV	$(x, y, \lambda) \underset{\delta}{\rightarrow} \lambda$	(for $x \in I$ and $y \in \{d, e\}$)
V	$(b, x, \lambda) \underset{\delta}{\rightarrow} x'$	(for $x \in B \cup I$)
VI	$(x, y, \lambda) \underset{\delta}{\rightarrow} \lambda$ $(x, y, \lambda) \underset{\delta}{\rightarrow} x$	(for $x \in B \cup I \cup \{b\}$ and $y \in C_1$) (for $x, y \in C_1$)
VII	$(x, y, \lambda) \underset{\delta}{\rightarrow} yx$	(for $x \in C_1$ and $y \in B \cup I$)

VIII	$\begin{cases}(x', f, \lambda) \underset{\delta}{\rightarrow} x'' \\ (b, f, \lambda) \underset{\delta}{\rightarrow} b''\end{cases}$	(for $x \in B \cup I$)
IX	$\begin{cases}(x, y, \lambda) \underset{\delta}{\rightarrow} \lambda \\ (x, y, \lambda) \underset{\delta}{\rightarrow} x\end{cases}$	(for $x \in B \cup I \cup \{b\}$ and $y \in C_2$) (for $x \in C_1$ and $y \in C_2$)
X	$(x'', x, \lambda) \underset{\delta}{\rightarrow} f$	(for $x \in B \cup I$)
XI	$(x'', y, \lambda) \underset{\delta}{\rightarrow} h$	(for $x \in B \cup I \cup \{b\}$ and $y \in B \cup I \cup \{a\}$ such that $x \neq y$ and $xy \neq ba$)
XII	$(b, a, \lambda) \underset{\delta}{\rightarrow} gcg$	

XIII All the other productions are of the form $(x, y, \lambda) \underset{\delta}{\rightarrow} y$.

It is easily seen that $A^*\{ba\}A^* \cap L(G) \neq \Phi$ iff our instance of the Post correspondence problem has a solution. The proof of the fact that there are $J\langle 0, 1\rangle L$ systems with the required properties is analogous. ▮

THEOREM 3.3. *Let $k \geqslant 2$ and $l \geqslant 0$ (resp. $k \geqslant 0$ and $l \geqslant 2$). Then $J\langle k, l\rangle\mathscr{L} \subset J\langle k-1, l+1\rangle\mathscr{L}$ (resp. $J\langle k, l\rangle\mathscr{L} \subset J\langle k+1, l-1\rangle\mathscr{L}$).*

Proof. We omit the proof since it is quite the same as that of Theorem 6.1 in Herman and Rozenberg (1974). ▮

Inductive use of Theorem 3.3 gives us the following corollary.

COROLLARY 3.1. *Let k, l, k' and l' be positive integers such that $k + l = k' + l'$. Then $J\langle k, l\rangle\mathscr{L} = J\langle k', l'\rangle\mathscr{L}$.* ▮

The following two theorems will be useful later on. For the sake of brevity we denote $A^k \times A \times A^l$ by $A_{k,l}$ for any alphabet A.

THEOREM 3.4. *Let $G = (A, \delta, g, \omega)$ be a $J\langle k, l\rangle L$ system. If there is a nonnegative integer N such that, for each $(P, Q) \in \Rightarrow_G \cap (L(G) \times L(G))$, productions of $\delta \cap (A_{k,l} \times A^+)$ can be applied to at most N symbols of P in the direct derivation $P \Rightarrow_G Q$, then $L(G)$ is finite.*

Proof. Let

$$M = \max\{\lg(P) \mid \delta \cap (A_G \times \{P\}) \neq \Phi\}.$$

It is easily seen that no word of $L(G)$ can have more than $\max\{(k + l + N + 2)M, \lg(\omega)\}$ symbols. Hence $L(G)$ is finite. ▮

THEOREM 3.5. *Each $J\langle k, l\rangle L$ language L can be generated by some $J\langle k', l'\rangle L$ system $G = (A, \delta, g, \omega)$ such that*

$$\delta \cap (A_G \times A_g{}^*\{g\}A_g{}^*) \subset A_{k,l} \times A_g{}^*,$$

for some nonnegative integers k' and l' such that $k' + l' = k + l$. Especially G can be chosen to be a $J\langle k, l\rangle L$ system if at least one of the five conditions below is satisfied

$$k = l = 0, \quad k + l = 1, \quad kl \geqslant 1, \quad \lambda \in L, \quad L \cap (A \cup \cdots \cup A^{\max\{k,l\}-1}) = \Phi. \tag{1}$$

Proof. Suppose $L = L(G')$ for some $J\langle k, l\rangle L$ system $G' = (A, \delta', g, \omega')$. We transform G' to another $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$. If $k = l = 0$, then $G = G'$. Assume then that $k + l \geqslant 1$. First $\delta \cap (A_{k,l} \times A_g{}^*) = \delta' \cap (A_{k,l} \times A_g{}^*)$. The remaining part of δ is defined by the following rules I–III.

I If $k \geqslant 2$, then, for all words Pa and Q such that

$$P \in B_G \setminus (A^k \cup \{g\}A^{k-1}),$$

$a \in A$ and $Q \in C_G \setminus \{g^l\}$ (this is for $l \geqslant 1$; $Q = \lambda$ for $l = 0$), define $(P, a, Q) \to_\delta \lambda$. Correspondingly, if $l \geqslant 2$, then define $(P, a, Q) \to_\delta \lambda$ for all words P and aQ such that $P \in A^k$, $a \in A$ and $Q \in C_G \setminus (A^l \cup A^{l-1}\{g\})$.

II If $k \geqslant 1$, then, for all words Pa, Q, R and S such that $P \in A^{k-1}$, $a \in A$, $Q \in C_G$, $R \in (A^*\{g\})^*$, $S \in A^*$ and Pa as an initial subword of PaQ directly G'-generates RS (this is for $Q \in A^l$; $PaQ \Rightarrow_{G'} RS$ for $Q \in C_G \setminus A^l$), define $(gP, a, Q) \to_\delta S$. Correspondingly, if $l \geqslant 1$, then, for all words P, aQ, R and S such that $P \in A^k$, $a \in A$, $Q \in A^{l-1}$, $R \in A^*$, $S \in (\{g\}A^*)^*$ and aQ as a final subword of PaQ directly G'-generates RS, define $(P, a, Qg) \to_\delta R$.

III If $k, l \geqslant 1$, then, for all words $P_i a$, Q_i and R_i such that $P_i \in A^i$, $a \in A$, $Q_i \in (A^*\{g\})^*$, $R_i \in A^*$ and $P_i a \Rightarrow_{G'} Q_i R_i$, define $(g^{k-i}P_i, a, g^l) \to_\delta R_i$ (for $i = 0, \ldots, k-1$).

Finally let $P_1, \ldots, P_n$ be exactly those words of $L(G')$ (if any, otherwise $\omega = \omega'$) which are directly G'-generated by some word of $A \cup \cdots \cup A^{k+l}$. Then we define $\omega = \omega' g P_1 \cdots g P_n$.

It is easily seen that G is of the desired form. It is also seen that $L(G) = L(G')$ if either $k = l = 0$ or $k + l = 1$ or $kl \geqslant 1$ or

$$L(G') \cap (A \cup \cdots \cup A^{\max\{k,l\}-1}) = \Phi$$

and that $L(G) = L(G') \cup \{\lambda\}$ if $k + l \geqslant 2$, $kl = 0$ and

$$L(G') \cap (A \cup \cdots \cup A^{\max\{k,l\}-1}) \neq \Phi.$$

If in the latter case $\lambda \notin L(G')$, then by Theorem 3.3 $L(G') = L(G'')$ for some $J\langle k', l'\rangle L$ system G'' such that $k' + l' = k + l$ and $k'l' \geqslant 1$, and we transform G'' in the way described above. Hence the theorem holds. ∎

It is seen by Lemma 3.1 (the bracketed claim) that the process described in the previous proof is not effective in general.

Theorem 3.5 gives us a normal form for $J\langle k, l\rangle L$ systems. In this normal form no fragmentation takes place in the leftmost k symbols and in the rightmost l symbols of a word (or in a word of length less than $\max\{k, l\}$). In the theorem we cannot always choose G to be a $J\langle k, l\rangle L$ system unless at least one of conditions (1) is fulfilled. This is seen by the following lemma.

LEMMA 3.2. *For each integer $k \geqslant 2$ (resp. $l \geqslant 2$) there is a $J\langle k, 0\rangle L$ language (resp. a $J\langle 0, l\rangle L$ language) which cannot be generated by any $J\langle k, 0\rangle L$ system (resp. $J\langle 0, l\rangle L$ system) $G = (A, \delta, g, \omega)$ such that*

$$\delta \cap (A_G \times A_g{}^*\{g\}A_g{}^*) \subset A_{k,0} \times A_g{}^*$$

(resp. $\delta \cap (A_G \times A_g{}^*\{g\}A_g{}^*) \subset A_{0,l} \times A_g{}^*$).

Proof. We prove only the unbracketed claim, the proof of the bracketed one being analogous. Let $k \geqslant 2$ and $L = L_1 \cup \{a^{k-1}\}$ where

$$L_1 = \{a^{2^n+2k} \mid n \geqslant 0\}.$$

Then L can be generated by the $DJ\langle k, 0\rangle L$ system $G' = (A, \delta', g, a^{1+2k})$ where $A = \{a\}$ and

$$\begin{cases} (a^k, a, \lambda) \underset{\delta'}{\to} a^2, \\ (ga^{k-1}, a, \lambda) \underset{\delta'}{\to} g, \\ (g^n a^{k-n}, a, \lambda) \underset{\delta'}{\to} a \qquad (\text{for } n = 2,\dots, k). \end{cases}$$

Suppose $L = L(G)$ for some $J\langle k, 0\rangle L$ system G as described in the claim. If $\delta \cap (\{(a^k, a, \lambda)\} \times \{a\}^+) = \Phi$, then by Theorem 3.4 $L(G)$ is finite which is not the case. So there is a production $(a^k, a, \lambda) \to_\delta a^i$ for some integer $i \geqslant 1$. It is easily seen that this production is deterministic (otherwise $L(G)$ contains words not in L) and also that the rest of the productions of G are deterministic. Hence G is a $DF\langle k, 0\rangle L$ system.

Let $a^k \Rightarrow_G a^r$ and

$$a^{2^{n'}+2k} \underset{G}{\Rightarrow} a^{2^n+2k}$$

for some large integers n' and n. Then we have the equation

$$i(2^{n'} + k) + r = 2^n + 2k.$$

This can hold only if $i = 2^u$, for some $u \geqslant 0$, and $ik + r = 2k$. Now, if $u = 0$, then $L(G)$ is finite. So $u \geqslant 1$. But then $r = 0$ and $a^{k-1} \Rightarrow_G \lambda$, a contradiction. ∎

Theorem 3.5 and its proof immediately yield the following corollary.

COROLLARY 3.2. *Let* $G = (A, \delta, g, \omega)$ *be a* $J\langle k, l\rangle L$ *system such that* $\delta \cap (A_{k,l} \times A_g{}^*\{g\}A_g{}^*) = \Phi$. *Then* $L(G) \in F\langle k', l'\rangle\mathscr{L}$ *for some nonnegative integers* k' *and* l' *such that* $k' + l' = k + l$. *If in addition at least one of conditions* (1) *in the statement of Theorem* 3.5 *is satisfied, then* $L(G) \in F\langle k, l\rangle\mathscr{L}$.

We give an alternative way of defining $J\langle k, l\rangle L$ languages. We say that, for a symbol q, a $\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ is *q-guarded* iff either $q \notin A$ or $q \in A$ and

(i) $\delta \cap ((B_G \times \{q\} \times C_G) \times A_g{}^*) = ((B_G \times \{q\} \times C_G) \times \{q\})$;

$(\text{ii})_1$ if $k \geqslant 1$, then, for each symbol $a \in A$ and all words $P_i \in (A \setminus \{q\})^i$, $Q_i \in \bigcup_{n=0}^{i} \{g^n\}A^{i-n}$, $R \in C_G$ and $S \in A^*$, the following equality holds true:

$$(Q_{k-i-1}qP_i\,, a, R) \underset{\delta}{\rightarrow} S \qquad \text{iff} \quad (g^{k-i}P_i\,, a, R) \underset{\delta}{\rightarrow} S$$
$$(\text{for } i = 0, \ldots, k-1);$$

$(\text{ii})_2$ if $l \geqslant 1$, then, for each symbol $a \in A$ and all words $P_i \in (A \setminus \{q\})^i$, $Q_i \in \bigcup_{n=0}^{i} A^{i-n}\{g^n\}$, $R \in B_G$ and $S \in A^*$, the following equality holds true:

$$(R, a, P_iqQ_{l-i-1}) \underset{\delta}{\rightarrow} S \qquad \text{iff} \quad (R, a, P_ig^{l-i}) \underset{\delta}{\rightarrow} S$$
$$(\text{for } i = 0, \ldots, l-1).$$

Thus if, for a q-guarded $\langle k, l\rangle L$ system G, q appears in some word of $(L(G)$ it acts as a permanent inactive (condition (i)) information blockage to the left (condition $(\text{ii})_1$) and to the right (condition $(\text{ii})_2$).

The following theorem establishes the connection between $J\langle k, l\rangle L$ systems and $\langle k, l\rangle L$ systems which are q-guarded for some symbol q.

THEOREM 3.6. *A language L is a $J\langle k, l\rangle L$ language iff there exists a symbol q and a q-guarded $\langle k, l\rangle L$ system G such that $L = J_q(L(G))$.*

Proof. (1) Let $G = (A, \delta, g, \omega)$ be a $J\langle k, l\rangle L$ system and q a symbol not in A_g. We transform G to a q-guarded $\langle k, l\rangle L$ system $G' = (A \cup \{q\}, \delta', g, \omega')$ such that $L(G) = J_q(L(G'))$, as follows. For any word $P \in A_g{}^*$, let $\bar{P}$ be the word which we get from P by replacing each g by q. Then $\omega' = \bar{\omega}$. For each production $(P, a, Q) \rightarrow_\delta R$ we define $(P, a, Q) \rightarrow_{\delta'} \bar{R}$. The rest of the productions are given by conditions (i), $(\text{ii})_1$ and $(\text{ii})_2$ above.

(2) Let $G = (A, \delta, g, \omega)$ be a q-guarded $\langle k, l\rangle L$ system for some symbol q. If $A = \{q\}$, then $J_q(L(G)) = L(G')$ for any $J\langle k, l\rangle L$ system $G' = (A', \delta', g, \omega')$ with $\omega' \in \{g\}^*$. Assume then that $A \neq \{q\}$. We transform G to a $J\langle k, l\rangle L$ system $G'' = (A \setminus \{q\}, \delta'', g, \omega'')$ such that $J_q(L(G)) = L(G'')$, as follows. For any word $P \in A^*$ let $\bar{P}$ be the word which we get from P by replacing each q by g. Then $\omega'' = \bar{\omega}$. Let

$$\bar{\delta} = \{(X, \bar{P}) \mid X \underset{\delta}{\rightarrow} P\}.$$

Then $\delta'' = \bar{\delta} \cap (B_{G''} \times A_g{}^*)$ where $B = A \setminus \{q\}$. ∎

Since conditions $(\text{ii})_1$ and $(\text{ii})_2$ are not needed for $k = l = 0$, this alternative way of defining $J\langle k, l\rangle L$ languages is especially convenient in the case of $J0L$ languages (cf. Rozenberg, Ruohonen and Salomaa (1974)).

We show next that removing any of the conditions (i), $(\text{ii})_1$ and $(\text{ii})_2$ in the definition of a q-guarded $\langle k, l\rangle L$ system causes Theorem 3.6 to fail.

LEMMA 3.3.

$$L_1 = \{a^n b a^n \mid n \geqslant 0\} \cup \{a^n b a^{2n} \mid n \geqslant 0\} \cup \{c\} \notin JI\mathscr{L},$$
$$L_2 = \{a^n b a^n \mid n \geqslant 0\} \cup \{a^{2n} b a^n \mid n \geqslant 0\} \cup \{c\} \notin JI\mathscr{L}$$

Proof. We prove only the former claim, the proof of the latter one being analogous. Let

$$L_3 = \{a^n b a^n \mid n \geqslant 0\}, \qquad L_4 = \{a^n b a^{2n} \mid n \geqslant 0\}$$

so that $L_1 = L_3 \cup L_4 \cup \{c\}$. Suppose L_1 can be generated by some $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ where $A = \{a, b, c\}$. Then

$$\delta \cap (\{(a^k, a, a^l)\} \times A_g{}^*) \subset A_G \times \{a, g\}^*$$

(each word of L_1 contains only one occurrence of b or c). By Theorem 3.4 we see that if $\delta \cap (\{(a^k, a, a^l)\} \times \{a\}^+) = \Phi$, then L_1 is finite which is not the case. So there is a production $(a^k, a, a^l) \to_\delta a^i$ for some $i \geqslant 1$. It is easy to see that this production is deterministic, otherwise $L(G)$ contains words not in L_1. Hence $\delta \cap (\{(P, x, Q)\} \times A_g{}^*) \subset \{(P, x, Q)\} \times A^*$ for any word

$$PxQ \in \{a^{k+l}b, a^{k+l-1}b, \ldots, ba^{k+l}\}.$$

By Corollary 3.2 we may now suppose that G is an $F\langle k, l\rangle L$ system.

It follows that for sufficiently long words P and Q of L_1, $\{(P, Q)\} \cap \Rightarrow_G \subset (L_3 \times L_3) \cup (L_4 \times L_4)$ and, in case $P \Rightarrow_G Q$, the productions used in the direct derivation are deterministic (otherwise again $L(G)$ contains words not in L_1).

Now let $a^{n'}ba^{n'} \Rightarrow_G a^nba^n$ for some $n, n' \geqslant k + l$. Then $a^{n'}ba^{2n'} \Rightarrow_G a^nba^{n+in'}$ and so $in' = n$. We see that $L(G)$ is finite if $i = 1$. Hence $i \geqslant 2$. But then

$$\{a^nba^n \mid n \geqslant 2(k + l)\}$$
$$= \bigcup_{t=1}^{T} \{a^{f_n}ba^{f_n} \mid f_n = (n_t - x)\, i^{n-n_t} + x,\ n \geqslant n_t\}$$

for some integers $T \geqslant 1$ and $n_1, \ldots, n_T \geqslant 2(k + l)$ and some rational number x. This cannot be the case and so we have a contradiction. ∎

Now $L_1 = J_q(L(G_1)) = J_q(L(G_2))$ and $L_2 = J_q(L(G_3))$ where $G_i = (A, \delta_i, g, c)$ (for $i = 1, 2, 3$), $A = \{a, b, c, q\}$ and the productions are

$$(\lambda, a, \lambda) \xrightarrow[\delta_1]{} a, \qquad (\lambda, b, \lambda) \xrightarrow[\delta_1]{} aba, \qquad (\lambda, c, \lambda) \xrightarrow[\delta_1]{} bqb,$$

$$(\lambda, q, \lambda) \xrightarrow[\delta_1]{} aq;$$

$$(\lambda, a, x) \xrightarrow[\delta_2]{} a, \qquad (x, a, \lambda) \xrightarrow[\delta_3]{} a \qquad (\text{for } x \in A_g \setminus \{q\});$$

$$(\lambda, a, q) \xrightarrow[\delta_2]{} a^2, \qquad (q, a, \lambda) \xrightarrow[\delta_3]{} a^2;$$

$$(\lambda, b, x) \xrightarrow[\delta_2]{} aba, \qquad (x, b, \lambda) \xrightarrow[\delta_3]{} aba \qquad (\text{for } x \in A_g \setminus \{q\});$$

$$(\lambda, b, q) \xrightarrow[\delta_2]{} aba^2, \qquad (q, b, \lambda) \xrightarrow[\delta_3]{} a^2ba;$$

$$(\lambda, c, x) \xrightarrow[\delta_2]{} bqb, \qquad (x, c, \lambda) \xrightarrow[\delta_3]{} bqb \qquad (\text{for } x \in A_g);$$

$$(\lambda, q, x) \xrightarrow[\delta_2]{} q, \qquad (x, q, \lambda) \xrightarrow[\delta_3]{} q \qquad (\text{for } x \in A_g).$$

Hence all the conditions (i), $(\text{ii})_1$ and $(\text{ii})_2$ in the definition of a q-guarded $\langle k, l\rangle L$ system are necessary for Theorem 3.6 to hold.

Since $L_1 = J_q(L(G_1))$ and G_1 is a $0L$ system, the following theorem holds.

THEOREM 3.7. *None of the families $\langle k, l\rangle\mathscr{L}$, $F\langle k, l\rangle\mathscr{L}$, $J\langle k, l\rangle\mathscr{L}$, $I\mathscr{L}$ and $JI\mathscr{L}$ is closed under the operator J_q.* ∎

We conclude this section by two theorems which together with Theorems 3.1 and 3.3 establish the hierarchy of the families $J\langle k, l\rangle\mathscr{L}$.

THEOREM 3.8. *Let k, l, k' and l' be nonnegative integers such that $k + l < k' + l'$. Then $\langle k', l'\rangle\mathscr{L} \setminus J\langle k, l\rangle\mathscr{L} \neq \Phi$.*

Proof. Let

$$L = \{a^{2^n+2(k'+l')} \mid n \geqslant 0\}.$$

Then L can be generated by the $D\langle k', l'\rangle L$ system $G' = (A, \delta', g, a^{1+2(k'+l')})$ where $A = \{a\}$ and

$$\begin{cases} (a^{k'}, a, a^{l'}) \underset{\delta'}{\rightarrow} a^2 \\ (P, a, Q) \underset{\delta'}{\rightarrow} \lambda \qquad \text{(for } P \in B_{G'} \text{ and } Q \in C_{G'} \text{ such that } PQ \neq a^{k'+l'}). \end{cases}$$

Suppose $L = L(G)$ for some $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ where $A = \{a\}$. By an argumentation similar to that in the beginning of the proof of the previous lemma we may suppose further that G is an $F\langle k, l\rangle L$ system. It is easily seen that then G must be deterministic. Hence $(a^k, a, a^l) \rightarrow_\delta a^i$ for some unique $i \geqslant 1$ (if $i = 0$, then $L(G)$ is finite by Theorem 3.4) and $a^{k+l} \Rightarrow_G a^r$ for some unique $r \geqslant 0$.

Now let

$$a^{2^{n'}+2(k'+l')} \underset{G}{\Rightarrow} a^{2^n+2(k'+l')}$$

for some large n' and n. Then we have the equation

$$i(2^{n'} + 2k' + 2l' - k - l) + r = 2^n + 2k' + 2l'.$$

This cannot hold unless $i = 1$ (and $n = n'$). But then $L(G)$ is finite, a contradiction. ∎

THEOREM 3.9.

$$\langle 1, 0\rangle\mathscr{L} \setminus J\langle 0, l\rangle\mathscr{L} \neq \Phi,$$
$$\langle 0, 1\rangle\mathscr{L} \setminus J\langle k, 0\rangle\mathscr{L} \neq \Phi$$

Proof. Let

$$L_1 = \{a^{2^n} \mid n \geqslant 0\} \cup \{ba^{2^n+1} \mid n \geqslant 0\} \cup \{c, \lambda\},$$

$$L_2 = \{a^{2^n} \mid n \geqslant 0\} \cup \{a^{2^n+1}b \mid n \geqslant 0\} \cup \{c, \lambda\}.$$

The proof of the fact that $L_1 \in \langle 1, 0\rangle \mathscr{L} \setminus J\langle 0, l\rangle \mathscr{L}$ and $L_2 \in \langle 0, 1\rangle \mathscr{L} \setminus J\langle k, 0\rangle \mathscr{L}$, following the main lines of the previous proof, is left to the reader (see also the proof of Theorem 6.7 in Herman and Rozenberg (1974)). ∎

The hierarchy of the families $J\langle k, l\rangle \mathscr{L}$ is summarized in the diagram of Fig. 1.

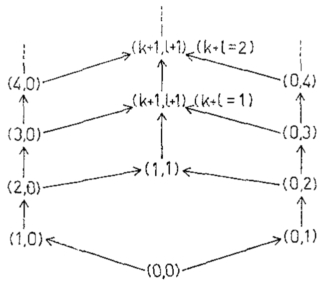

FIG. 1. The hierarchy of the families $J\langle k, l\rangle \mathscr{L}$. A family $J\langle k, l\rangle \mathscr{L}$ is denoted simply by (k, l). An arrow denotes strict inclusion. Two families are incomparable (though not disjoint since $J\langle 0, 0\rangle \mathscr{L} \neq \Phi$) unless they are connected by a directed path.

4. ERASING FRAGMENTATION

A $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ is called a *$J\langle k, l\rangle L$ system with erasing fragmentation* (or a $j\langle k, l\rangle L$ system in short) iff $\delta \subset A_G \times (A^* \cup \{g\}^*)$. Thus, whenever a symbol in some word P of A^+ causes fragmentation, its contribution to words that are directly G-generated by P is empty.

It is immediately verified that Theorems 3.1, 3.2, 3.4, 3.7, 3.8 and 3.9 are valid when the letters J are replaced by j's. Also Theorem 3.3 holds true for $j\langle k, l\rangle L$ languages. Since the method of proof of Theorem 6.1 in Herman and Rozenberg (1974) is not readily applicable in this case, we give a proof of this fact.

THEOREM 4. *Let $k \geqslant 2$ and $l \geqslant 0$ (resp. $k \geqslant 0$ and $l \geqslant 2$). Then $j\langle k, l\rangle \mathscr{L} \subset j\langle k-1, l+1\rangle \mathscr{L}$ (resp. $j\langle k, l\rangle \mathscr{L} \subset j\langle k+1, l-1\rangle \mathscr{L}$).*

Proof. We prove only the unbracketed claim, the proof of the bracketed one being analogous. Let $G = (A, \delta, g, \omega)$ be a $j\langle k, l\rangle L$ system for some $k \geqslant 2$ and $l \geqslant 0$. We transform G to a $j\langle k-1, l+1\rangle L$ system $G' = (A, \delta', g, \omega')$ such that $L(G') = L(G)$, as follows. The production relation δ' is given by the following rules I–IV.

I For any production $(g^k, a, g^l) \rightarrow_\delta P$, define $(g^{k-1}, a, g^{l+1}) \rightarrow_{\delta'} P$.

II For any word $P \in B_{G'} \setminus \{g^{k-1}\}$ and any symbol $a \in A$, define $(P, a, g^{l+1}) \rightarrow_{\delta'} \lambda$.

III For any words Pa, bQ and R such that $Pa \in B_G \setminus \{g^{k-1}\} A, a \in A$, $b \in A$, $Q \in C_G$, $R \in A_g{}^*$ and $(Pa, b, Q) \rightarrow_\delta R$, define $(P, a, bQ) \rightarrow_{\delta'} R$.

IV For any words a, bQP, R and S such that $a \in A$, $b \in A$, $QP \in C_G$, $P \in A_g$ (this is for $l \geqslant 1$; $P = \lambda$ for $l = 0$), $(g^{k-1}a, b, QP) \rightarrow_\delta R$ and $(g^k, a, bQ) \rightarrow_\delta S$ (this is for $l \geqslant 1$; $(g^k, a, \lambda) \rightarrow_\delta S$ for $l = 0$), define

$$(g^{k-1}, a, bQP) \underset{\delta'}{\rightarrow} \begin{cases} \lambda, & \text{if } R \in \{g\}^+ \\ R, & \text{if } R \notin \{g\}^+ \quad \text{and} \quad S \in \{g\}^+ \\ SR, & \text{if } R,\ S \notin \{g\}^+. \end{cases}$$

Finally let $P_1, \ldots, P_n$ be exactly those words of $L(G)$ (if any, otherwise $\omega' = \omega$) which can be G-generated directly by some word of $A^2 \cup \cdots \cup A^{l+2}$. Then define $\omega' = \omega g P_1 \cdots g P_n$.

It is easily seen that $L(G') = L(G)$. ∎

$j\langle 2, 0\rangle L$ systems similar to the $J\langle 1, 0\rangle L$ systems G' in Lemma 3.1 can be used to show that the process described in the previous proof is not effective in general (contrasting with the process in the proof of Theorem 6.1 in Herman and Rozenberg (1974)). The question, whether or not this process (and also any of the processes in the proofs of Theorems 3.2 and 3.5) can be replaced by an effective one, remains unanswered within this paper. We conjecture however that the answer is negative.

From Theorem 4.1 it now follows that Theorem 3.5 is valid for $j\langle k, l\rangle L$ languages. (Note also that the same is true for Lemma 3.2 and the discussion preceding it.) It also follows that the diagram in Fig. 1 holds true when the letters J are replaced by j's. It is interesting to note that this diagram becomes somewhat more complicated if only nonfragmented axioms are used in the generating $j\langle k, l\rangle L$ systems (the languages thus obtained are called $\bar{F}j\langle k, l\rangle L$ languages; see Fig. 2). Further discussion of such systems lies outside the scope of this paper.

The following theorem establishes the position of the family $jI\mathscr{L}$ with respect to $JI\mathscr{L}$ and $I\mathscr{L}$.

THEOREM 4.2.

(i) $J0\mathscr{L} \setminus jI\mathscr{L} \neq \Phi$,

(ii) $j0\mathscr{L} \setminus I\mathscr{L} \neq \Phi$.

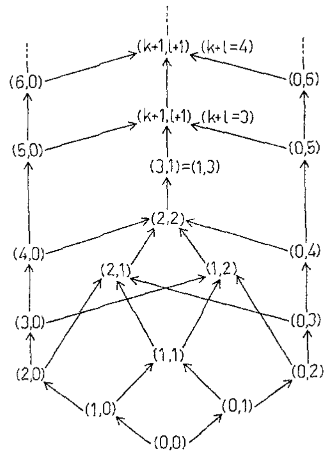

FIG. 2. The hierarchy of the families $\bar{F}j\langle k, l\rangle\mathscr{L}$. A family $\bar{F}j\langle k, l\rangle\mathscr{L}$ is denoted by (k, l). Otherwise the notation is the same as in Fig. 1.

Proof. (i) Let $L = L_1 \cup L_2 \cup L_3$ where

$$\begin{aligned} L_1 &= \{a^n b a^{2n} c a^{2n} b a^n \mid n \geqslant 0\}, \\ L_2 &= \{a^{n+m} b a^{2n+m+1} \mid n \geqslant 0, m \geqslant 0\}, \\ L_3 &= \{a^{2n+m+1} b a^{n+m} \mid n \geqslant 0, m \geqslant 0\}. \end{aligned}$$

Then L can be generated by the $J0L$ system $G' = (A, \delta', g, bcbgabgba)$ where $A = \{a, b, c\}$ and δ' is given by

$$\begin{cases} (\lambda, a, \lambda) \underset{\delta'}{\rightarrow} a \\ (\lambda, b, \lambda) \underset{\delta'}{\rightarrow} aba \\ (\lambda, c, \lambda) \underset{\delta'}{\rightarrow} aca \\ (\lambda, c, \lambda) \underset{\delta'}{\rightarrow} a^2ga^2. \end{cases}$$

Suppose now that $L = L(G)$ for some $j\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$. Since Theorem 3.5 holds for $j\langle k, l\rangle L$ systems, we may assume that

$$\delta \cap (A_G \times A_g{}^*\{g\}A_g{}^*) \subset A_{k,l} \times A_g{}^*.$$

Moreover $\lambda \notin L$ so that there are no productions $(a^k, a, a^l) \rightarrow_\delta g^j$ for $j \geqslant 1$. Also $\delta \cap (\{(a^k, a, a^l)\} \times (A^* \setminus \{a\}^*)) = \Phi$ (each word of L contains at most three symbols differing from a). Hence $(a^k, a, a^l) \rightarrow_\delta a^i$ for some integer

$i \geqslant 1$ (if $(a^k, a, a^l) \to_\delta \lambda$ were the only production of the form, then by Theorem 3.4 $L(G)$ would be finite).

It follows that, for sufficiently long words $P \in L_1$, $\Rightarrow_G \cap (L \times \{P\}) \subset L_1 \times L_1$. So, for each $(Q, x, R) \in A_G$ such that QxR is a subword of $g^k P g^l$ where P is a sufficiently long word of L_1, there exists a word $S \in A^*$ such that $(Q, x, R) \to_\delta S$. It is easily seen that for $QxR \notin \{ca^{k+l}, aca^{k+l-1}, \ldots, a^{k+l}c\}$ these productions are deterministic (the assumption that $(Q, x, R) \to_\delta g$ for some $QxR \in \{ba^{k+l}, aba^{k+l-1}, \ldots, a^{k+l}b\}$ implies that $L(G)$ contains words not in L). Thus each sufficiently long word of L_1 G-generates directly exactly one longer word of L_1 (and possibly some words of $L_2 \cup L_3$). By an argument similar to that finishing the proof of Lemma 3.3 we see now that $(a^k, a, a^l) \to_\delta a$.

Obviously sufficiently long words of $L_2 \cup L_3$ can be directly G-generated by some word of L_1 only through fragmentation. Now suppose $(Q, x, R) \to_\delta g$ for some $QxR \in \{ca^{k+l}, aca^{k+l-1}, \ldots, a^{k+l}c\}$. Then, for some large n' and $n + m$,

$$a^{n'}ba^{2n'}ca^{2n'}ba^{n'} \underset{G}{\Rightarrow} a^{n+m}ba^{2n+m+1}.$$

So also

$$a^{n'}ba^{2n'}ca^{2n'}ba^{n'} \underset{G}{\Rightarrow} a^{n+m}ba^{2(n+m)}ca^{2(n+m)}ba^{n+m}$$

and hence $2(n + m) \geqslant 2n + m + 1$ i.e., $m \geqslant 1$. Thus it is seen that sufficiently long words $a^n ba^{2n+1}$ cannot be directly G-generated by any word of L_1. (Note that in a word $a^{n+m}ba^{2n+m+1} \in L_2$ n and m are uniquely determined.)

Suppose next that

$$a^{2n'+m'+1}ba^{n'+m'} \underset{G}{\Rightarrow} a^n ba^{2n+1}$$

for some large $n' + m'$ and n. But then

$$a^{2(n'+n+1)+m'+1}ba^{(n'+n+1)+m'} \underset{G}{\Rightarrow} a^{3n+2}ba^{3n+2},$$

a contradiction. Finally suppose that

$$a^{n'+m'}ba^{2n'+m'+1} \underset{G}{\Rightarrow} a^n ba^{2n+1}$$

for some large $n' + m'$ and n. Then $n > n' + m' \geqslant n'$ and

$$a^{2(n-n'-1)+p+1}ba^{(n-n'-1)+p} \underset{G}{\Rightarrow} a^{p+3(n-n')-m'-1}ba^{p+3(n-n')-m'-1}$$

for $p \geqslant k + l + m'$, again a contradiction.

(ii) Let $L = L_1 \cup L_2 \cup L_3$ where

$$L_1 = \{a^n ba^{2n} ca^{2n} ba^n \mid n \geqslant 1\},$$
$$L_2 = \{a^{n+m} ba^{2n+m} \mid n \geqslant 1, m \geqslant 1\},$$
$$L_3 = \{a^{2n+m} ba^{n+m} \mid n \geqslant 1, m \geqslant 1\}.$$

The proof of the fact that $L \in jI\mathscr{L} \setminus I\mathscr{L}$ goes along the main lines of the proof of part (i) (see also the proof of Theorem 12 in Rozenberg, Ruohonen and Salomaa, 1974). It is left to the reader. ∎

As an immediate consequence of the previous theorem (and Theorem 7.2 in Herman and Rozenberg, 1974), the following chain of strict inclusions holds true (the first one is a well-known result)

$$\langle k, l\rangle\mathscr{L} \subsetneq F\langle k, l\rangle\mathscr{L} \subsetneq j\langle k, l\rangle\mathscr{L} \subsetneq J\langle k, l\rangle\mathscr{L}.$$

5. Some Nonclosure Results

The main result of this section is that the family $JI\mathscr{L}$ is an anti-*AFL* i.e., it is not closed under any of the *AFL*-operations (union, λ-free homomorphism, intersection with regular languages, inverse homomorphism and λ-free catenation closure).

First we need the following lemma.

Lemma 5.1. *Let* $L = L_1 \cup L_2$ *where*

$$L_1 = \{a^{2^n} ba^{2^m} \mid n, m \geqslant 0\},$$
$$L_2 = \{a^{2^n} cdea^{2^m} \mid n, m \geqslant 0\}.$$

Then $L^+ \notin JI\mathscr{L}$.

Proof. Suppose L^+ can be generated by some $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ where $A = \{a, b, c, d, e\}$. Let

$$K_n = \{P \mid P \in L, \lg(P) \leqslant n\} \qquad (\text{for } n \geqslant 3)$$

and

$$H_n = \{P \mid P \in L^+ \text{ and the maximum number of consecutive } a\text{'s in } P \text{ is } n\} \qquad (\text{for } n \geqslant 1).$$

Now suppose that there exists a number $N \geqslant 3$ such that, for arbitrarily large values of n, there are words $Q_n \in K_N{}^+$ and $P_n \in H_n$ such that $Q_n \Rightarrow_G P_n$.

Then, for some sufficiently large $n = 2^p + 2^q$, there is a word $R \in L^{k+l+1}$ such that $Q_n = S_n R S_n' R S_n''$, for some S_n, S_n', $S_n'' \in A^*$, and both $RS_n'R$ and S_n' as subwords of Q_n directly, but not through fragmentation, G-generate some words of $\{a\}^+$. But then, for some $s \geqslant 1$,

$$S_n R(S_n'R)^j S_n'' \underset{G}{\Rightarrow} R_j$$

for every $j \geqslant 1$ and some $R_j \in H_{n+(j-1)s}$. This is a contradiction since some such $H_{n+(j-1)s}$'s are empty.

Hence each word of H_n can be directly G-generated only by some word of $H_{n'}$ such that n' grows along with n. There also must be a deterministic production $(a^k, a, a^l) \to_\delta a^i$ for some integer $i \geqslant 1$.

Consider a word

$$P_{n,m} = Ra^{2^n+2^m}R' \in H_{2^n+2^m}$$

for some $R, R' \in A^*$ and n, m sufficiently large. Let Ra^k and a^lR' as subwords of $P_{n,m}$ directly G-generate words Sa^r and a^sS' respectively, for some $S \in A^+\{b, cde\} \cup \{\lambda\}$, $S' \in \{b, cde\}A^+ \cup \{\lambda\}$ and $r, s \geqslant 0$, in such a way that

$$P_{n,m} \underset{G}{\Rightarrow} Sa^{i(2^n+2^m-k-l)+r+s}S'.$$

As is easily seen by considering all possible values of n and m, it must be the case that $r + s - ik - il = 0$ and $i = 2^u$ for some $u \geqslant 0$. If $u \geqslant 1$, then words of H_{2^n+1}, for sufficiently large n, are not directly G-generated by any word of L. On the other hand, if $u = 0$, then $\Rightarrow_G \cap (L \times H_n) \subset H_n \times H_n$ for large values of n. So we end up in a contradiction. ▌

THEOREM 5.1. *The families $JI\mathscr{L}$ and $jI\mathscr{L}$ are anti-AFL's.*

Proof. The theorem is proved by the following facts I–V.

I The language

$$L_1 = \{a^nba^n \mid n \geqslant 0\} \cup \{a^nba^{2n} \mid n \geqslant 0\} \cup \{c\} \notin JI\mathscr{L}$$

(Lemma 3.3) is a union of two $0L$ languages

$$\{a^nba^n \mid n \geqslant 0\} \qquad \text{and} \qquad \{a^nba^{2n} \mid n \geqslant 0\} \cup \{c\}.$$

II The language $L_1 = h_1(L_2)$ where

$$L_2 = \{a^nba^n \mid n \geqslant 0\} \cup \{a^nda^{2n} \mid n \geqslant 0\} \cup \{c\}$$

is a $0L$ language and h_1 is the λ-free homomorphism given by

$$h_1(a) = a, \qquad h_1(b) = b, \qquad h_1(c) = c, \qquad h_1(d) = b.$$

III The language $L_1 = L_3 \cap L_4$ where

$$L_3 = \{a^n d a^n \mid n \geqslant 0\} \cup \{a^n e a^{2n} \mid n \geqslant 0\} \cup L_1$$

is a $0L$ language and $L_4 = \{a, b, c\}^*$ is a regular language.

IV The language $L_1 = h_2^{-1}(L_3)$ where h_2 is the homomorphism given by

$$h_2(a) = a, \qquad h_2(b) = b, \qquad h_2(c) = c, \qquad h_2(d) = d^2, \qquad h_2(e) = e^2.$$

V The language $L_5 = L_6{}^+$ where

$$L_6 = \{a^{2^{2n}} b a^{2^{2m}} \mid n, m \geqslant 0\} \cup \{a^{2^{2n}} cdea^{2^{2m}} \mid n, m \geqslant 0\}$$

is not a JIL language (Lemma 5.1). However $L_6 = L(G)$ for the $j0L$ system $G = (A, \delta, g, abagacdea)$ where $A = \{a, b, c, d, e\}$ and δ is given by

$$(\lambda, a, \lambda) \underset{\delta}{\rightarrow} a^2, \qquad (\lambda, b, \lambda) \underset{\delta}{\rightarrow} b, \qquad (\lambda, b, \lambda) \underset{\delta}{\rightarrow} cde,$$

$$(\lambda, c, \lambda) \underset{\delta}{\rightarrow} ba, \qquad (\lambda, d, \lambda) \underset{\delta}{\rightarrow} g, \qquad (\lambda, e, \lambda) \underset{\delta}{\rightarrow} ab. \quad \blacksquare$$

It follows from the previous theorem and its proof that *each of the families* $J\langle k, l\rangle\mathscr{L}$ *and* $j\langle k, l\rangle\mathscr{L}$ *is an anti-AFL*, too.

6. Comparison with Other Families of Languages

We discuss the position of the families $JI\mathscr{L}$, $J0\mathscr{L}$, $jI\mathscr{L}$ and $j0\mathscr{L}$ with respect to the families $0\mathscr{L}$, $F0\mathscr{L}$, $T0\mathscr{L}$, $FT0\mathscr{L}$, $JT0\mathscr{L}$, $E0\mathscr{L}$, $ET0\mathscr{L}$ and $I\mathscr{L}$ of developmental languages and also to the families of regular languages $\mathscr{L}(RG)$, context free languages $\mathscr{L}(CF)$, context sensitive languages $\mathscr{L}(CS)$ and recursively enumerable languages $\mathscr{L}(RE)$ of the Chomsky hierarchy.

THEOREM 6.1. $T0\mathscr{L} \setminus JI\mathscr{L} \neq \Phi$.

Proof. Consider the language

$$L = \{a^{2^n \cdot 3^m} \mid n, m \geqslant 0\}.$$

It is known that $L \in T0\mathscr{L} \setminus I\mathscr{L}$ (see the proof of Theorem 10.9 in Herman and Rozenberg (1974)). Now suppose that $L = L(G)$ for some $J\langle k, l\rangle L$ system $G = (A, \delta, g, \omega)$ where $A = \{a\}$. As in the proofs of Lemma 3.3 and Theorem 3.8 we may assume that G is an $F\langle k, l\rangle L$ system and so L is an $F\langle k, l\rangle L$ language. But then (see Theorem 7.2 in Herman and Rozenberg (1974)) L is also an IL language, a contradiction. $\blacksquare$

THEOREM 6.2.

(i) $\mathscr{L}(RG) \setminus JI\mathscr{L} = \{\Phi\}$,

(ii) $\{\Phi\} \subsetneq \mathscr{L}(CF) \setminus JI\mathscr{L}$,

(iii) $\{\Phi\} \subsetneq \mathscr{L}(RG) \setminus JT0\mathscr{L}$.

Proof. (i) It is well-known (see Theorem 4 in Rozenberg and Lee (1973)) that $\mathscr{L}(RG) \setminus I\mathscr{L} = \{\Phi\}$. Since by definition $\Phi \notin JI\mathscr{L}$, (i) follows.

(ii) The language

$$\{a^n b a^n \mid n \geqslant 0\} \cup \{a^n b a^{2n} \mid n \geqslant 0\} \cup \{c\} \notin JI\mathscr{L}$$

(Lemma 3.3) is easily seen to be a context free language.

(iii) It is shown in Rozenberg, Ruohonen and Salomaa (1974) (Theorem 4) that the regular language $\{a^{2n+1} \mid n \geqslant 1\} \cup \{a^2\}$ is not a $JT0L$ language. ▮

THEOREM 6.3. $j0\mathscr{L} \setminus TF0\mathscr{L} \neq \Phi$.

Proof. Let

$$L = \{aba^n \mid n \geqslant 1\} \cup \{a^n ba \mid n \geqslant 1\} \cup \{acdca^n \mid n \geqslant 1\} \cup \{a^n cdca \mid n \geqslant 1\}.$$

Then L is generated by the $Dj0L$ system (A, δ, g, aba) where $A = \{a, b, c, d\}$ and

$$\begin{cases} (\lambda, a, \lambda) \underset{\delta}{\rightarrow} a, \\ (\lambda, b, \lambda) \underset{\delta}{\rightarrow} cdc, \\ (\lambda, c, \lambda) \underset{\delta}{\rightarrow} aba, \\ (\lambda, d, \lambda) \underset{\delta}{\rightarrow} g. \end{cases}$$

It is easily seen that L is not a $TF0L$ language (see the proof of Theorem 11 in Rozenberg, Ruohonen and Salomaa, 1974). ▮

An *extended* $J\langle k, l\rangle L$ *system* (an $EJ\langle k, l\rangle L$ system in short) is an ordered pair $G = (\bar{G}, \Delta)$ where $\bar{G} = (A, \delta, g, \omega)$ is a $J\langle k, l\rangle L$ system and $\Delta \subset A$. The extension of a $j\langle k, l\rangle L$ system (an $Ej\langle k, l\rangle L$ system) is defined analogously. The language generated by an $EJ\langle k, l\rangle L$ system $G = (\bar{G}, \Delta)$ is defined by

$$L(G) = L(\bar{G}) \cap \Delta^*.$$

In an obvious way we can now define the families $EJI\mathscr{L}$, $EjI\mathscr{L}$, $EPJI\mathscr{L}$ and $EPjI\mathscr{L}$ which all coincide with the family of recursively enumerable languages, as is seen by the following theorem.

THEOREM 6.4. $EJI\mathscr{L} = EPI\mathscr{L} = EjI\mathscr{L} = EPjI\mathscr{L} = \mathscr{L}(RE)$.

Proof. It is well-known that $E\langle 1, 0\rangle\mathscr{L} = \mathscr{L}(RE)$ (see Theorem 7.3 in Herman and Rozenberg, 1974). Hence $EJI\mathscr{L} = EjI\mathscr{L} = \mathscr{L}(RE)$.

Let L be a recursively enumerable language. Then $L = L(G)$ for some $E\langle 1, 0\rangle L$ system $G = (\bar{G}, \Delta) = ((A, \delta, g, \omega), \Delta)$. We transform G to an $EPj\langle 1, 0\rangle L$ system $G' = ((A', \delta', g, \omega g\omega\#), \Delta)$ such that $L(G') = L(G)$, as follows. The alphabet A' is the union of the disjoint alphabets A and

$$B_1 = \{[P] \mid P \in A_g A\},$$
$$B_2 = \{\#, \S, \%, \&\},$$
$$B_3 = \{(P) \mid P \in (A \cup \{\%, \S\})^2(A \cup \{\%, \S, \&\})\}$$

and δ' is given by the following schemes I–XVI.

I	$(g, x, \lambda) \to_{\delta'} [gx]$	(for $x \in A$)
II	$([xy], z, \lambda) \to_{\delta'} [yz]$	(for $x \in A_g$ and $y, z \in A$)
III	$(x, [yz], \lambda) \to_{\delta'} \S P$	(for $x, y \in A_g$, $z \in A$ and $P \in A^*$ such that $(y, z, \lambda) \to_\delta P$)
IV	$(x, \#, \lambda) \to_{\delta'} \&\#$	(for $x \in B_1$)
V	$(x, \#, \lambda) \to_{\delta'} \&$	(for $x \in B_1$)
VI	$(x, \S, \lambda) \to_{\delta'} \%$	(for $x \in \{g, \%\}$)
VII	$(\%, \&, \lambda) \to_{\delta'} g$	
VIII	$(\%, x, \lambda) \to_{\delta'} (\%\%x)$	(for $x \in A$)
IX	$((xyz), u, \lambda) \to_{\delta'} (yzu)$	(for $x, y \in A \cup \{\%, \S\}$, $z \in A \cup \{\S\}$ and $u \in A \cup \{\S, \&\}$)
X	$(x, (yzu), \lambda) \to_{\delta'} z$	(for $x, y \in A \cup \{\%, \S\}$, $z \in A \cup \{\%\}$ and $u \in A$)
XI	$(x, (y\S z), \lambda) \to_{\delta'} y$	(for $x \in A \cup \{\%, \S\}$ and $y, z \in A \cup \{\S\}$)
XII	$(x, (y\S\&), \lambda) \to_{\delta'} \&y$	(for $x \in A \cup \{\%, \S\}$ and $y \in A$)
XIII	$(x, (\S\S\&), \lambda) \to_{\delta'} \&$	(for $x \in A \cup \{\S\}$)
XIV	$(x, (yz\S), \lambda) \to_{\delta'} \S$	(for $x, y \in A \cup \{\%, \S\}$ and $z \in A$)
XV	$(x, (yz\&), \lambda) \to_{\delta'} \&z$	(for $x, y \in A \cup \{\%, \S\}$ and $z \in A$)
XVI	All the other productions are of the form $(x, y, \lambda) \to_\delta y$.	

The idea of the above construction is roughly the following. A symbol to be rewritten as $P \in A^*$ is rewritten as $\S P$. So we get a propagating system. The §'s are then removed to the beginning of the scanned word where they are cut away. (Cf. Theorem 9.9 in Salomaa (1973) and the well-known fact that $EPI\mathscr{L} = \mathscr{L}(CS)$.) ▮

By Theorems 3.8, 4.2, and 6.1–6.4, Section 10 in Herman and Rozenberg (1974) and Theorem 11 in Rozenberg, Ruohonen and Salomaa (1974) the diagram in Fig. 3 holds true.

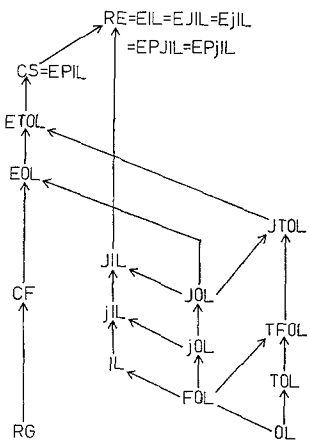

FIG. 3. The position of $J0\mathscr{L}$, $j0\mathscr{L}$, $JI\mathscr{L}$ and $jI\mathscr{L}$ with respect to some other families of languages. A family $X\mathscr{L}$ (resp. $\mathscr{L}(X)$) is denoted by XL (resp. X). Otherwise the notation is the same as in Fig. 1. No two families are disjoint since $\mathscr{L}(RG) \cap 0\mathscr{L} \neq \Phi$.

RECEIVED: November 4, 1974

REFERENCES

HERMAN, G. AND ROZENBERG, G. (1974), "Developmental Systems and Languages," North-Holland, Amsterdam.

LINDENMAYER, A. (1968), Mathematical models for cellular interactions in development I, II, *J. Theor. Biol.* **18**, 280.

ROZENBERG, G. AND LEE, K. P. (1973), Some properties of the class of L-languages with interactions, submitted for publication, available as technical report.

ROZENBERG, G., RUOHONEN, K., AND SALOMAA, A. (1974), Developmental systems with fragmentation, *Internat. J. Comput. Math.*, to appear.

SALOMAA, A. (1973), "Formal Languages," Academic Press, New York.